

PERINGKASAN DOKUMEN TEKS OTOMATIS BERDASARKAN SEBUAH KUERI MENGUNAKAN *BIDIRECTIONAL LONG SHORT TERM MEMORY NETWORK*

AUTOMATIC TEXT DOCUMENT SUMMARY BASED ON A QUERY USING BIDIRECTIONAL LONG SHORT TERM MEMORY NETWORK

Yuliska¹, Khairul Umam Syaliman²

^{1,2}Politeknik Caltex Riau
yuliska@pcr.ac.id¹

ABSTRACT

Query-focused multi-document extractive summarization aims to produce a single, short document that summarizes a set of long documents that are relevant to a given query in an extractive manner. During the past few years, deep learning approaches have been utilized to generate summaries in an abstractive or extractive manner. In this study, we utilized Bidirectional long short term memory network (Bi-LSTM) to produce a concise summary based on a single query from a set of long text documents by employing extractive approach. Bi-LSTM is one of popular deep learning methods that is considered good for text classification. We used DUC 2005-2007 benchmark datasets in our experiment. Our Experiment shows that Bi-LSTM can produce relatively good summaries with the score of ROUGE-1 = 43.53, ROUGE-2 = 11.40 and ROUGE-L = 18.67.

Keywords: *Query-Focused Summarization, Extractive Summarization, Bi-LSTM*

ABSTRAK

Query-focused summarization atau peringkasan teks otomatis berdasarkan sebuah kueri adalah sebuah bidang penelitian pada natural language processing yang bertujuan untuk menghasilkan sebuah dokumen pendek atau ringkasan dari sekumpulan dokumen panjang, dimana ringkasan yang dihasilkan harus relevan dengan sebuah kueri yang diberikan. Hingga saat ini, berbagai metode deep learning telah digunakan untuk menghasilkan ringkasan dari sebuah maupun banyak dokumen dengan pendekatan abstraktif maupun ekstraktif. Pada penelitian ini, peneliti menggunakan Bidirectional Long Short Term Memory Network (Bi-LSTM) untuk menghasilkan sebuah ringkasan berdasarkan sebuah kueri dari beberapa dokumen dengan pendekatan ekstraktif. Bi-LSTM merupakan salah satu metode deep learning yang sering digunakan dalam klasifikasi teks. Dataset yang peneliti gunakan adalah DUC 2005-2007 dataset, yang merupakan dataset yang umum digunakan pada text summarization. Berdasarkan eksperimen yang peneliti lakukan, Bi-LSTM mampu menghasilkan ringkasan yang baik, yang dibuktikan dengan skor ROUGE-1 = 43.53, skor ROUGE-2 = 11.40 dan skor ROUGE-L = 18.67.

Kata Kunci: Peringkasan Teks, Kueri, Extractive Summarization, Bi-LSTM

PENDAHULUAN

Automatic Text summarization atau peringkasan teks secara otomatis baru-baru ini mendapat banyak perhatian dalam bidang *natural language processing* karena penerapannya yang sangat menjanjikan di berbagai bidang. Contoh penerapannya yaitu generasi *snippet* pada mesin pencari, generasi headline berita, *question answering (QA)*, generasi *Minutes of Meeting (MoM)*, dan mesin rekomendasi yang dipersonalisasi. Diantara penerapan-penerapan tersebut, *query-focused summarization (QFS)*, yaitu sebuah *task* yang bertujuan untuk menghasilkan

ringkasan dari sekumpulan dokumen panjang berdasarkan sebuah kueri, hingga saat ini masih menjadi penelitian yang hangat.

Baru-baru ini, metode-metode *deep learning* telah diterapkan dalam *automatic text summarization*, dan sebagian besar pendekatan yang digunakan adalah bersifat ekstraktif (Kobayashi et al., 2015) (Ren et al., 2018) (Cao et al., 2016). Pendekatan ekstraktif masih lebih disukai untuk peringkasan teks karena pendekatan ini dinilai tidak begitu kompleks dibandingkan dengan pendekatan abstraktif.

Penelitian tentang *query-focused extractive summarization* sejauh ini telah menggunakan berbagai macam metode dan pendekatan. Pada awal kemunculannya, *query-focused extractive summarization* sebagian besar menggunakan metode *unsupervised* berbasis graf untuk mengekstrak kalimat penting dan relevan terhadap kueri inputan (Srivanthi et al., 2008) (Bhaskar & Bandyopadhyay, 2010). Pada pendekatan ini, node adalah kalimat dan *edge score* (nilai tepi) adalah kesamaan antar kalimat. Setiap node diberi bobot relevansi berdasarkan relevansinya dengan kueri. Lalu, pendekatan *supervised machine learning* diterapkan untuk meringkas dokumen-dokumen panjang berdasarkan sebuah kueri. (Ouyang et al., 2011), (Daumé & Marcu, 2006), (Conroy et al., 2005) menggunakan *Support Vector Regression (SVR)*, *Bayesian Statistical Model (BAYESUM)*, dan *Hidden Markov Model (HMM)*, untuk mengekstraks fitur yang bergantung pada kueri maupun fitur yang tidak ada hubungannya dengan kueri. Kemudian, dilakukan sebuah perhitungan untuk mengestimasi tingkat kepentingan kalimat.

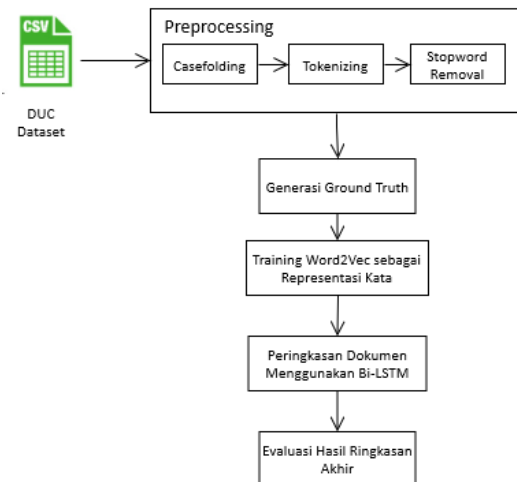
Dengan semakin populernya metode *deep learning* dalam bidang klasifikasi teks, semakin banyak pula sistem peringkasan otomatis (*automatic text summarization system*) yang juga menerapkan metode *deep learning* untuk menangani *automatic text summarization* maupun *query-focused summarization*. (Vig et al., 2022) memperkenalkan dua buah metode dalam *query-focused summarization*, yang disebut *RELREG (Relevance Regression)* dan *SEGENC (Segment Encoder)*. *RELREG* adalah sebuah model yang di-*training* untuk memprediksi *ROUGE* yang berisikan antara dokumen sumber dan ringkasan referensi. Sedangkan *SEGENC* adalah dokumen sumber dipecah ke dalam beberapa segmen, dimana kemudian segmen ini akan ditambahkan ke kueri secara terpisah. *RELREG* dan *SEGENC* menggunakan *encoder* berbasis *neural network*. (Baumel Tal, 2018) kemudian menggunakan

pendekatan *seq2seq model* untuk menghasilkan ringkasan berdasarkan sebuah kueri. *Seq2seq model* yang digunakan yaitu *Recurrent Neural Network (RNN)* yang ditambahkan dengan *attention mechanism*.

Pada penelitian ini, peneliti menggunakan metode *Bidirectional LSTM* untuk mengekstrak kalimat-kalimat penting dari banyak dokumen panjang, dimana kalimat yang dihasilkan diharapkan memiliki relevansi yang tinggi dengan kueri inputan. *Bidirectional LSTM* digunakan karena metode ini merupakan salah satu metode yang populer dalam klasifikasi teks. Pada penelitian ini, peneliti menggunakan pendekatan klasifikasi teks dalam mengekstrak kalimat penting, sehingga *Bidirectional LSTM* merupakan metode yang paling tepat untuk diterapkan.

METODE

Bab ini menjelaskan metodologi penelitian yang berupa tahapan-tahapan yang dilalui untuk mencapai tujuan penelitian.



Gambar 1. Metodologi Penelitian

Data Yang Digunakan

Data yang digunakan dalam eksperimen pada penelitian ini adalah dataset *DUC 2005-2007*. Dataset ini terdiri dari dokumen teks yang berupa artikel berita berbahasa Inggris yang dikelompokkan dalam beberapa topik. Pada masing-masing topik, terdapat empat ringkasan referensi. Terdapat 145 topik, dengan total kalimat adalah 104.773

kalimat. Pada eksperimen yang peneliti lakukan, *DUC 2005* digunakan sebagai data training, *DUC 2006* sebagai data validation, dan *DUC 2007* sebagai data testing. Contoh mengenai data yang digunakan dapat dilihat pada tabel 1 sebagai berikut:

Tabel 1. Data Yang Digunakan

Tahun	Jumlah Topic	Jumlah Kalimat	Sumber Artikel
2005	50	45931	TREC
2006	50	34560	AQUAINT
2007	45	24282	AQUAINT

Berikut adalah contoh data yang digunakan pada penelitian ini:

DUC	cluster_id	sent_id	sentence_no_stopwords	sentence_with_stopwords	question
2007	0	0	mobile ala ap white supremacist arrested buyin...	mobile ala ap a white supremacist arrested aft...	activities morris dees southern poverty law ce...
2007	0	1	chris scott gilliam said want like unabomber k...	chris scott gilliam said he did not want to be...	activities morris dees southern poverty law ce...
2007	0	2	wanted kill everybody david pasqualotto specia...	he wanted to kill everybody david pasqualotto ...	activities morris dees southern poverty law ce...

Gambar 2. Contoh Data yang Digunakan

Preprocessing

Preprocessing yang dilakukan pada penelitian ini terdiri dari 3 tahapan *preprocessing*, yaitu:

- 1) *Casefolding*: adalah tahapan mengubah huruf besar menjadi huruf kecil. Contohnya adalah “*GraDuatEs*” menjadi “*graduates*”
- 2) *Tokenizing*: adalah tahapan dimana pemecahan kalimat menjadi kumpulan kata-kata. Contohnya adalah sebagai berikut:
 [aside from that only 63 percent of indians are high school graduates]
 Menjadi:
 [‘aside’, ‘from’, ‘that’, ‘only’, ‘63’, ‘percent’, ‘of’, ‘indians’, ‘are’, ‘high’, ‘school’, ‘graduates’]
- 3) *Stopword removal*: adalah tahapan mengeliminasi kata-kata yang yang tidak penting. Contohnya adalah sebagai berikut:
 [‘aside’, ‘from’, ‘that’, ‘only’, ‘63’, ‘percent’, ‘of’, ‘indians’, ‘are’, ‘high’, ‘school’, ‘graduates’]

Menjadi:

[‘aside’, ‘only’, ‘63’, ‘percent’, ‘indians’, ‘high’, ‘school’, ‘graduates’]

Generasi Ground Truth

Untuk proses training, dibutuhkan kalimat yang telah dilabeli (*ground truth*). *Ground truth* dalam penelitian ini adalah dalam bentuk pelabelan biner (0/1) yang diberikan kepada setiap kalimat. Namun, *DUC* sebagai *benchmark dataset* dalam *text summarization* hanya berisi ringkasan abstrak yang dibuat oleh manusia secara manual sebagai ringkasan referensi. Untuk mengatasi masalah ini, peneliti menggunakan pendekatan *unsupervised* untuk mengkonversi ringkasan referensi ke dalam bentuk kalimat yang dilabeli. Secara khusus, diberikan kumpulan kalimat $S = [s_1, s_2, s_3, \dots, s_n]$ dan kumpulan ringkasan referensi $M = [m_1, m_2, \dots, m_n]$, peneliti memberikan skor *ROUGE (Recall-Oriented Understudy for Gisting Evaluation)* (Lin, 2004) untuk setiap kalimat dengan cara membandingkan kalimat dengan ringkasan referensi. *ROUGE* yang digunakan adalah 4 varian *ROUGE*; *ROUGE-1*, *ROUGE-2*, *ROUGE-L* dan Skor *ROUGE-SU*. *ROUGE* adalah *evaluation metric* dalam peringkasan teks, sekaligus merupakan salah satu *similarity measures*. Berikut merupakan persamaan *ROUGE* yang dipakai dalam penelitian ini:

$$ROUGE - N = \frac{Count_{match}(Referensi_{N-gram}, Hasil Ringkasan_{N-gram})}{Count(Referensi_{N-gram})}$$

$$ROUGE - L = \frac{LCS(Hasil Ringkasan_{kata}, Referensi_{kata})}{Count(Referensi_{kata})}$$

$$ROUGE - SU = \frac{Count_{match}(SU(Referensi, Hasil Ringkasan))}{Count(SU(Referensi))}$$

Untuk *ROUGE-N*, penelitian ini menggunakan nilai $N=1,2$. *LCS* adalah *Longest Common Subsequence*, yaitu *subsequence* terpanjang yang mirip antara ringkasan referensi dan ringkasan yang

dihasilkan sistem atau aplikasi. *SU* adalah *skip-bigram* dengan tambahan unigram sebagai unit penghitung. $Count_{match}$ pada *ROUGE-N* dan *ROUGE-SU* adalah jumlah dari *N-gram* dan *SU* yang mirip antara ringkasan referensi dan ringkasan yang dihasilkan sistem atau aplikasi. $Count$ adalah jumlah *N-gram*, kata dan *SU* pada *ROUGE-N*, *ROUGE-L* dan *ROUGE-SU*. Kemudian, peneliti mengurutkan kalimat secara *descending* berdasarkan skor *ROUGE* kalimat tersebut. Selanjutnya, label “1” diberikan kepada K kalimat dengan *ROUGE* skor tertinggi, dan label “0” kepada kalimat sisa. Nilai K dalam hal ini adalah 10% dari jumlah total kalimat. Pendekatan ini diambil karena kalimat dengan label “1” adalah kalimat yang akan memberikan *ROUGE* skor maksimal. Sehingga, dari hasil *training*, akan dihasilkan model yang memberikan *membership score* tinggi terhadap kalimat dengan label “1”.

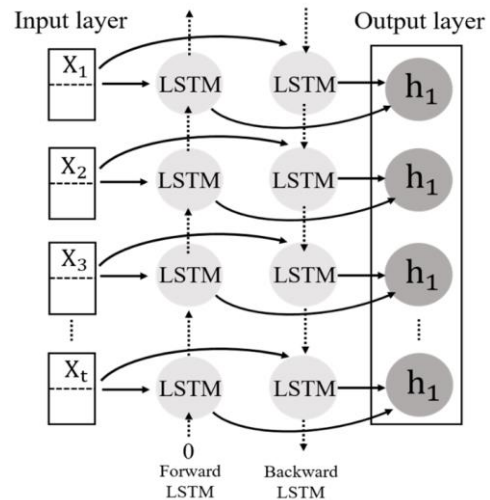
Word2Vec

Word2Vec (Mikolov et al., 2013) adalah salah satu cara merepresentasikan teks dalam bentuk *dense vector* yang disebut juga dengan *word embedding*. *Word2Vec* dapat memiliki 50 sampai dengan 300 dimensi. *Word2Vec* mulai ramai digunakan dalam bidang *natural language processing* di tahun 2013, karena *Word2Vec* merupakan *dense vectors* yang dapat merepresentasikan hubungan antar kata dengan lebih baik (dibandingkan dengan *TF-IDF*). Google telah menyediakan *pretrained Word2Vec* yang dapat langsung digunakan tanpa perlu melatih atau *training* model *Word2Vec* untuk mendapatkan *word embedding*. Namun, pada penelitian ini dilakukan proses *fine-tuning* untuk menghasilkan *Word2Vec* yang lebih sesuai dengan konteks.

Bi-LSTM

Adalah sebuah *network* yang merupakan gabungan dari *forward* dan *backward LSTM*. Dibandingkan *LSTM*, *Bi-*

LSTM dapat menghasilkan representasi kalimat yang lebih baik karena *network* ini mengkombinasikan representasi kalimat secara *forward* maupun *backward*. Pada penelitian ini, peneliti juga menerapkan operasi *Max-pooling* yaitu operasi yang memilih nilai maksimum dari setiap dimensi di setiap *hidden unit*.



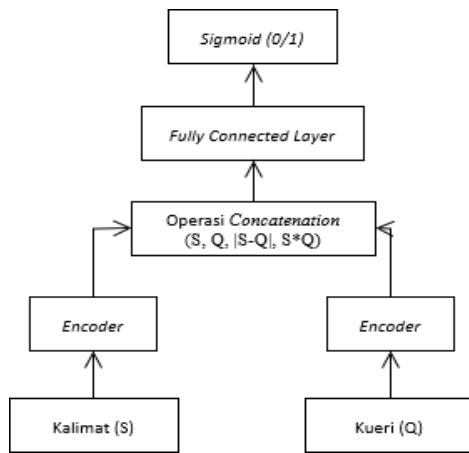
Gambar 3. Struktur *Bi-LSTM* (Jang et al., 2020)

Gambar 3 menjelaskan skema tentang bagaimana peneliti melakukan *training* terhadap model *deep learning* yang peneliti gunakan untuk menggenerasi ringkasan.

Dapat dilihat pada gambar 3 bahwa pada penelitian ini, model *deep learning* dianggap sebagai sebuah *encoder* yang dapat menggenerasi representasi kalimat maupun representasi kueri. Setelah representasi kalimat dan kueri didapatkan, dilakukan beberapa operasi berikut:

- 1) *Concatenation* (S, Q)
- 2) *Element-Wise product* ($S*Q$)
- 3) *Absolute element-wise difference* ($|S - Q|$)

Hasil dari operasi-operasi di atas adalah representasi yang menyatakan hubungan antara kueri dan kalimat. Representasi dari 3 operasi ini kemudian menjadi input ke *Fully Connected Network* yang bertugas mengklasifikasi kalimat (anggota ringkasan akhir atau tidak).



Gambar 4. Skema Training

Pemilihan Kalimat Ringkasan

Untuk memilih kalimat ringkasan, peneliti menggunakan *greedy approach* seperti penelitian-penelitian sebelumnya. Dimana pada saat *testing*, peneliti mengurutkan kalimat secara ascending menurut skor keanggotaan kalimat terhadap ringkasan akhir, yang didapatkan dari model. Peneliti menambahkan kalimat baru ke ringkasan akhir jika kalimat tersebut memiliki setidaknya 50% *bigram* baru. Hal ini akan terus dilakukan sampai ringkasan berisi lebih dan kurang 250 kata.

Evaluasi Ringkasan Akhir

Evaluasi kualitas ringkasan akhir dilakukan dengan cara membandingkan ringkasan akhir yang dihasilkan oleh *Bi-LSTM* dengan ringkasan referensi dari dataset. Perbandingan dilakukan dengan menggunakan persamaan *ROUGE-1*, *ROUGE-2*, dan *ROUGE-L*.

HASIL DAN PEMBAHASAN

Pada bagian ini dijelaskan mengenai skema training yang peneliti terapkan dalam membuat model *Bi-LSTM* untuk memilih kalimat penting dan juga penjelasan dari hasil eksperimen.

Skema Training

Sebelum melakukan *training*, *library NLTK* peneliti gunakan untuk melakukan *tokenizing* dan membersihkan kalimat. Sebagai perbandingan, penelitian ini juga melakukan eksperimen generasi *summary*

dengan menggunakan *network* gabungan *CNN* dan *LSTM* (*CNN+LSTM*). *Bi-LSTM*, *CNN* dan *LSTM* diimplementasikan dengan menggunakan *library Keras*. Kemudian, *pretrained Word2Vec* dengan jumlah dimensi 300 digunakan sebagai *word embedding* atau representasi kalimat. *Pretrained Word2vec* kemudian peneliti lakukan *fine-tuning* agar representasi kata menjadi lebih sesuai konteks. *Hidden size* yang peneliti gunakan pada model *Bi-LSTM* adalah 50, dengan *activation function* yang penulis gunakan yaitu *Tanh* untuk *Bi-LSTM* dan *LSTM*, dan *ReLU* pada *CNN*. Peneliti juga menerapkan operasi *Max-Pooling* pada model *Bi-LSTM* dengan ukuran *pooling* senilai 2. Operasi *dropout* juga diterapkan dengan rasio 0.5. Pada *MLP network* (*fully connected network*), peneliti menggunakan *hidden size* 100, 50, dan 1. *Optimizer* yang peneliti gunakan yaitu *adam* dengan *learning rate* senilai 10^{-2} dan *batch size* senilai 100.

Hasil Pengujian

Hasil pengujian dapat dilihat pada table 2, yaitu berupa skor *ROUGE* dari ringkasan akhir yang dihasilkan oleh 2 buah jenis model, yaitu *CNN+LSTM* dan *Bi-LSTM*. Semakin besar nilai *ROUGE*, maka semakin baik pula kualitas ringkasan akhir. Pada table 2, dapat dilihat bahwa *Bi-LSTM* memiliki performa yang lebih baik dibandingkan dengan *CNN+LSTM*. *Bi-LSTM* berhasil menggenerasi ringkasan akhir dengan nilai *ROUGE-1* sebesar 43.53, nilai *ROUGE-2* sebesar 11.40 dan *ROUGE-L* sebesar 18.67.

Tabel 2. Nilai *ROUGE* Ringkasan Akhir

Model	<i>ROUGE-1</i>	<i>ROUGE-2</i>	<i>ROUGE-L</i>
<i>CNN+LSTM</i>	41.89	10.24	17.93
<i>Bi-LSTM</i>	43.53	11.40	18.67

Pada table 3 di bawah, peneliti membandingkan hasil ringkasan akhir yang dihasilkan oleh *CNN+LSTM* dan *Bi-LSTM*. Peneliti menebalkan kata atau kalimat yang beririsan dengan ringkasan referensi (ringkasan yang dibuat oleh manusia). Dapat dilihat bahwa *Bi-LSTM* mampu menggenerasi kalimat yang lebih relevan dengan kueri. Sementara itu, ringkasan

yang dihasilkan oleh *CNN+LSTM* hanya mampu menggenerasi ringkasan yang memiliki irisan sangat sedikit dengan ringkasan akhir.

Tabel 3. Contoh Ringkasan Akhir dari 2 Model

Model	Kueri	Ringkasan Akhir dari Model	Ringkasan Referensi
<i>CNN+LSTM</i>	What countries are having chronic potable water shortage s and why?	The scarcity of water in the city would be partially eased with the completion of the world bank aided 175 million US dollar, Sayedabad water treatment project the sources added. Cyprus has also planned to encourage people to use water saving devices and systems for recycling grey water for use in toilets and garden irrigation. Hanoi, October 15th, a drought in Vietnam has caused a serious water shortage affecting about 3 million people.	Potable water shortages are common in many countries of Asia and Africa and have also occurred on the island of Cyprus and in Mexico. Damaged or obsolete distribution systems present problems in Angola, Bangladesh, China, Cyprus and Lebanon. The biggest problem is providing limited resources of water to growing urban areas where the need is the greatest and problems of waste and pollution are the most severe.
<i>Bi-LSTM</i>	Write an account of the sequence of events involving the Kennedy family during and following the plane crash that killed John F. Kennedy, Jr., his wife, and his sister-in-law.	After Sen. Edward Kennedy traveled to Bridgehampton, NY to meet with and console John Kennedy sister Caroline Kennedy Schlossberg, the Kennedy family released a statement. New York, July 17, Xinhua, the united states coast guard and the air national guard are conducting a massive search off the coast of long island, NY,	John F. Kennedy, Jr. took off from Fairfield, NJ at 8:38 p.m. Friday, July 16, 1999 with his wife, Carolyn Bessette Kennedy and sister-in-law Lauren Bessette on board, headed to Martha's Vineyard to drop off Lauren before continuing to Hyannis. Sen. Edward Kennedy

for a small plane carrying John F. Kennedy, Jr. son of the 35th united states president united states media reported Saturday. In a private closely guarded ceremony aboard a united states navy destroyer, the remains of Kennedy, 38, his wife Carolyn Bessette Kennedy, 33, and Lauren Bessette, her 34-year-old sister, were laid in the choppy Atlantic Ocean.	traveled to Bridgehampton, NY to console Kennedy's sister Caroline Kennedy Schlossberg, after which the family released a statement. The aircraft wreckage and bodies were located on Wednesday.
---	--

SIMPULAN

Pada penelitian ini dilakukan sebuah eksperimen peringkasan otomatis dari banyak dokumen berdasarkan sebuah kueri. Peneliti melakukan generasi ringkasan dengan menggunakan *Bi-LSTM* dan juga gabungan *network CNN+LSTM* sebagai perbandingan. Dari hasil pengujian didapatkan bahwa *Bi-LSTM* dapat menghasilkan ringkasan akhir berdasarkan sebuah kueri dengan sangat baik yang dibuktikan dengan nilai *ROUGE* yang relatif tinggi.

DAFTAR PUSTAKA

- Baumel Tal. (2018). *Query Focused Summarization Using Seq2seq Models* [Ben Gurion University of the Negev]. <https://www.cs.bgu.ac.il/~elhadad/nlpproj/pub/tal-phd.pdf>
- Bhaskar, P., & Bandyopadhyay, S. (2010). A Query Focused Multi Document Automatic Summarization. *Proceeding of Twenty-Four Pacific Asia Conference on Language, Information and Computation*, 545–554.
- Cao, Z., Li, W., Li, S., Wei, F., & Li, Y. (2016). AttSum: Joint Learning of

- Focusing and Summarization with Neural Attention. *Proceeding of COLING 2016, the 26th International Conference on Computational Linguistic: Technical Paper*, 547–556.
- Conroy, J. M., Schlesinger, J. D., & Goldstein Stewart, J. (2005). CLASSY Query-Based Multi-Document Summarization. *Proceeding of the Document Understanding Conf. Wksp. 2005 at the Human Language Technology Conf./Conf. on Empirical Methods in Natural Language Processing (HLT/EMNLP)*.
- Daumé, H., & Marcu, D. (2006). Bayesian Query-Focused Summarization. *Proceeding of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the ACL*, 305–312.
- Jang, B., Kim, M., Harerimana, G., Kang, S. U., & Kim, J. W. (2020). Bi-LSTM model to increase accuracy in text classification: Combining word2vec CNN and attention mechanism. *Applied Sciences (Switzerland)*, 10(17).
<https://doi.org/10.3390/app10175841>
- Kobayashi, H., Noguchi, M., & Yatsuka, T. (2015). Summarization Based on Embedding Distributions. *Proceeding of the 2015 Conference on Empirical Methods in Natural Language Processing*, 1984–1989.
- Lin, C.-Y. (2004). ROUGE: A Package for Automatic Evaluation of Summaries. *Proceedings of the ACL-04 Workshop*, 74–51.
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *1st International Conference on Learning Representations, ICLR 2013 - Workshop Track Proceedings*, 1–12.
- Ouyang, Y., Li, W., Li, S., & Lu, Q. (2011). Applying regression models to query-focused multi-document summarization. *Information Processing and Management*, 47(2), 227–237.
<https://doi.org/10.1016/j.ipm.2010.03.005>
- Ren, P., Chen, Z., Ren, Z., Wei, F., Nie, L., Ma, J., & de Rijke, M. (2018). Sentence relations for extractive summarization with deep neural networks. *ACM Transactions on Information Systems*, 36(4).
<https://doi.org/10.1145/3200864>
- Sravanthi, M., Chowdary, C. R., & Sreenivasa Kumar, P. (2008). QueSTS: A Query Specific Text Summarization System. *Proceeding of the Twenty- First International FLAIRS Conference*, 219–224.
- Vig, J., Fabbri, A. R., Kryści, W., Nski, K., Wu, C.-S., & Liu, W. (2022). Exploring Neural Models for Query-Focused Summarization. *Findings of the Association for Computational Linguistics: NAACL 2022*, 1455–1468